

Laboratorio 1 de Regresion Aplicada

Edgar Acuna

Febrero 17 del 2018

Leyendo los datos del archivo mortality.txt que estan en academic.uprm.edu/eacuna

```
muertes=read.table("http://academic.uprm.edu/eacuna/mortalidad.txt",header=T)
```

```
#Mostrando los datos
```

```
muertes
```

```
##          nacion porc.inmuniz tasa.mort
## 1          Bolivia          77        118
## 2           Brazil          69         65
## 3         Cambodia          32        184
## 4           Canada          85          8
## 5            China          94         43
## 6 Czech_Republic          99         12
## 7            Egypt          89         55
## 8         Ethiopia          13        208
## 9           Finland          95          7
## 10          France          95          9
## 11          Greece          54          9
## 12           India          89        124
## 13           Italy          95         10
## 14           Japan          87          6
## 15          Mexico          91         33
## 16          Poland          98         16
## 17 Russian_Federation          73         32
## 18          Senegal          47        145
## 19          Turkey          76         87
## 20   United_Kingdom          90          9
```

```
#Haciendo un plot de tasa de mortalidad versus porcentaje de inmunizacion
```

```
x=muertes$porc.inmuniz
```

```
y=muertes$tasa.mort
```

```
win.graph()
```

```
plot(x,y,xlab="porcentaje de inmunizacion", ylab="tasa de mortalidad")
```

```
title("regresion de tasa de mortalidad versus porcentaje de inmunizacion")
```

```
pais=muertes$nacion
```

```
text(x,y,labels=as.character(pais),cex=.65,col="red",srt=30)
```

```
#Haciendo el ajuste por minimos cuadrados
```

```
l1<-lsfit(x,y)
```

```
#Mostrando los resultados del ajuste minimo cuadratico
```

```
l1
```

```
## $coefficients
```

```
## Intercept      X
```

```
## 224.316289 -2.135869
```

```
##
```

```

## $residuals
## [1] 58.1456523 -11.9413026 28.0315310 -34.7673929 19.4554314
## [6] -0.8652218 20.7760846 11.4500132 -14.4086993 -12.4086993
## [11] -99.9793430 89.7760846 -11.4086993 -32.4956541 3.0478233
## [16] 0.9989088 -36.3978252 21.0695714 25.0097829 -23.0880461
##
## $intercept
## [1] TRUE
##
## $qr
## $qt
## [1] -263.856021 -220.220457 26.585255 -46.947099 5.453068
## [6] -15.880173 7.786309 13.851571 -28.613581 -26.613581
## [11] -105.881005 76.786309 -25.613581 -45.080395 -10.346988
## [16] -13.813525 -46.147320 16.585532 14.652735 -36.280340
##
## $qr
##      Intercept      X
## [1,] -4.4721360 -346.14332292
## [2,] 0.2236068 103.10577093
## [3,] 0.2236068 0.43961557
## [4,] 0.2236068 -0.07441967
## [5,] 0.2236068 -0.16170867
## [6,] 0.2236068 -0.21020257
## [7,] 0.2236068 -0.11321478
## [8,] 0.2236068 0.62389236
## [9,] 0.2236068 -0.17140745
## [10,] 0.2236068 -0.17140745
## [11,] 0.2236068 0.22624245
## [12,] 0.2236068 -0.11321478
## [13,] 0.2236068 -0.17140745
## [14,] 0.2236068 -0.09381723
## [15,] 0.2236068 -0.13261234
## [16,] 0.2236068 -0.20050379
## [17,] 0.2236068 0.04196567
## [18,] 0.2236068 0.29413390
## [19,] 0.2236068 0.01286933
## [20,] 0.2236068 -0.12291356
##
## $qraux
## [1] 1.223607 1.080761
##
## $rank
## [1] 2
##
## $pivot
## [1] 1 2
##
## $tol
## [1] 1e-07
##
## attr(,"class")
## [1] "qr"

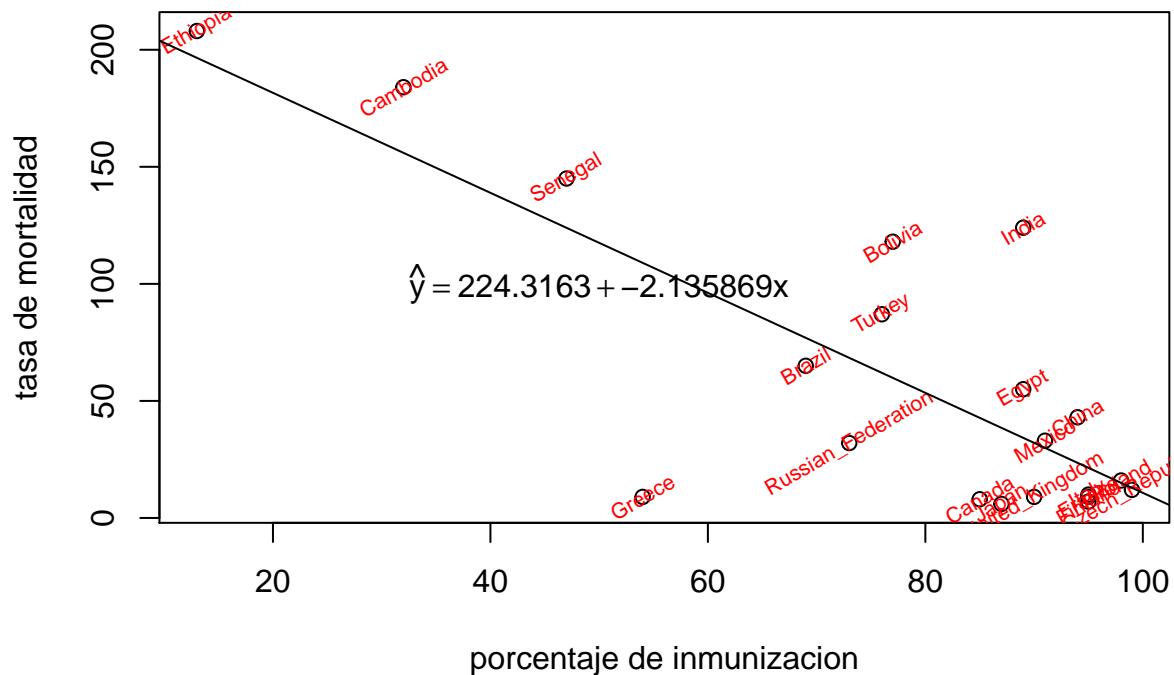
```

```
#Imprimiendo un resultado mas corto del ajuste minimocuadratico
ls.print(l1)
```

```
## Residual Standard Error=40.1393
## R-Square=0.6258
## F-statistic (df=1, 18)=30.1006
## p-value=0
##
##           Estimate Std.Err t-value Pr(>|t|)
## Intercept 224.3163 31.4403  7.1347      0
## X         -2.1359  0.3893 -5.4864      0
```

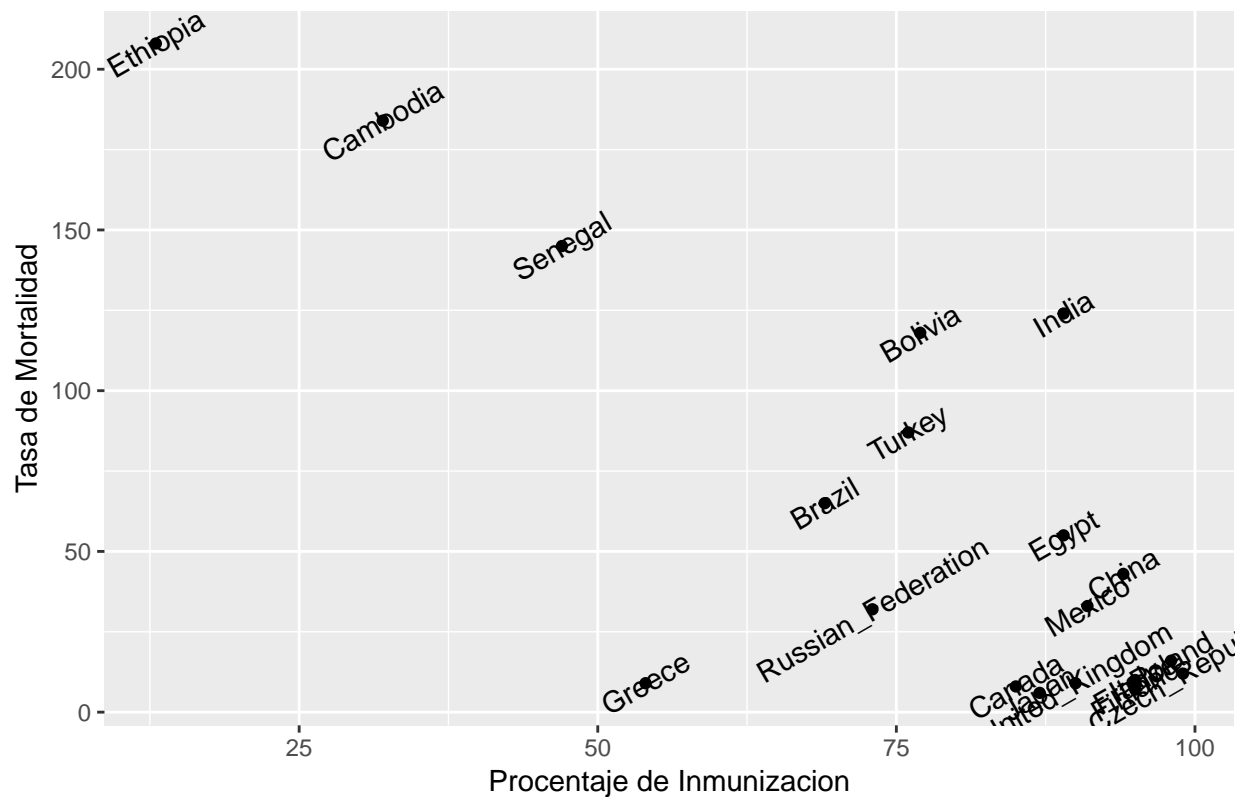
```
#Trazando la linea de regresión sobre el plot de puntos
abline(l1)
alfa=l1$coeff[1]
beta=l1$coeff[2]
text(50,100,bquote(hat(y)==.(alfa)+.(beta)*x))
#Ploteando la linea y los puntos usando ggplot2
library(ggplot2)
```

regresion de tasa de mortalidad versus porcentaje de inmunizacion



```
attach(muertes)
p=ggplot(muertes,aes(x=porc.inmuniz,y=tasa.mort,label=nacion))+geom_text(angle=30,size=4)+geom_point()
p+ggtitle("Relacion de Tasa de Mortalidad con Porcentaje de Inmunizacion") + xlab("Procentaje de Inmuniz")
```

Relacion de Tasa de Mortalidad con Porcentaje de Inmunizacion



```
# Calculo de la linea de regresion usando el comando lm
l3<-lm(tasa.mort~porc.immuniz,data=muerres)
l3

##
## Call:
## lm(formula = tasa.mort ~ porc.immuniz, data = muerres)
##
## Coefficients:
## (Intercept)  porc.immuniz
##      224.316      -2.136

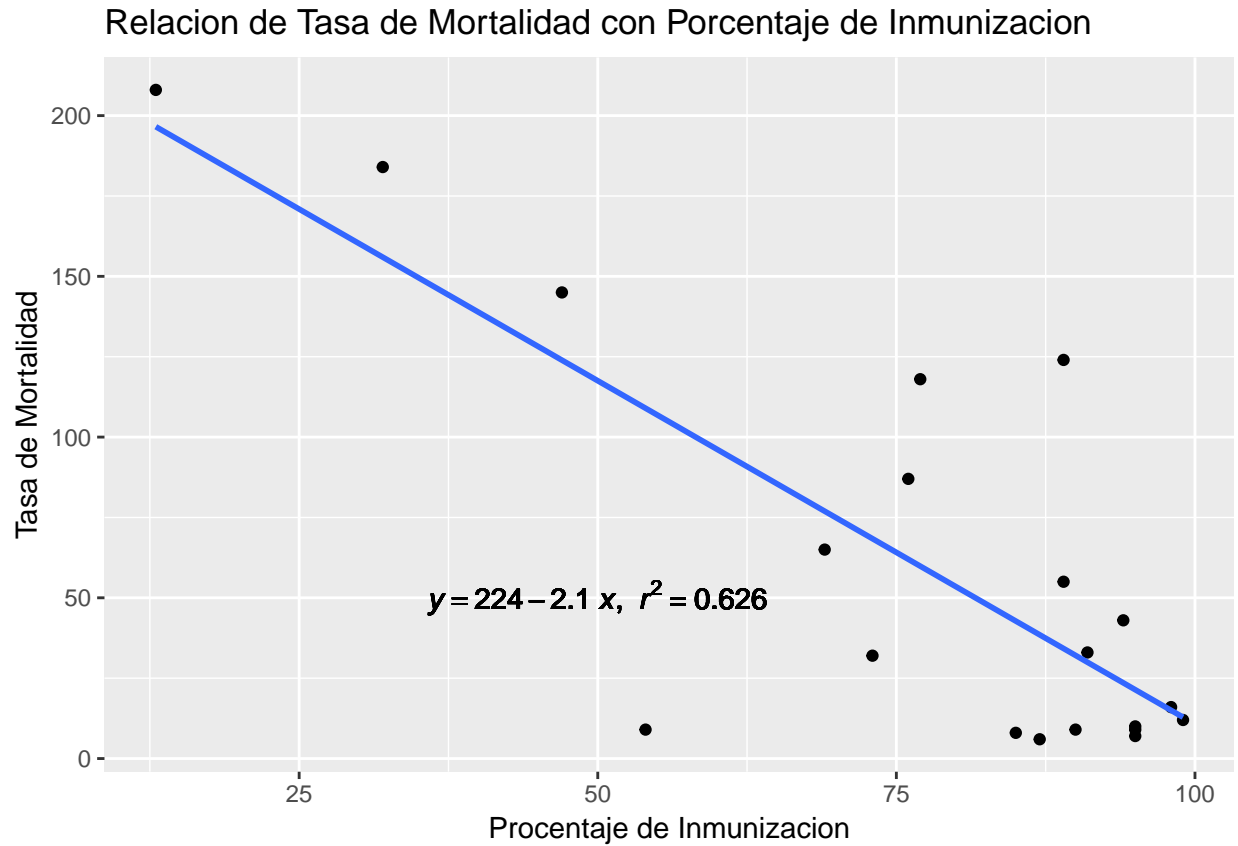
# Construimos una funcion para que nos imprima la ecuacion de la linea de regresion y el r2
lm_eqn = function(m) {

  l <- list(a = format(coef(m)[1], digits = 2),
           b = format(abs(coef(m)[2]), digits = 2),
           r2 = format(summary(m)$r.squared, digits = 3));

  if (coef(m)[2] >= 0) {
    eq <- substitute(italic(y) == a + b ~italic(x)*", "~italic(r)^2~"=="~r2,1)
  } else {
    eq <- substitute(italic(y) == a - b ~italic(x)*", "~italic(r)^2~"=="~r2,1)
  }

  as.character(as.expression(eq));
}
```

```
p=ggplot(muertes,aes(x=porc.inmuniz,y=tasa.mort))+geom_point()+geom_text(aes(x = 50, y = 50, label = lm.p))+ggtitle("Relacion de Tasa de Mortalidad con Porcentaje de Inmunizacion") + xlab("Procentaje de Inmuniz")
```



Hallando la prediccion para un valor dado de X.

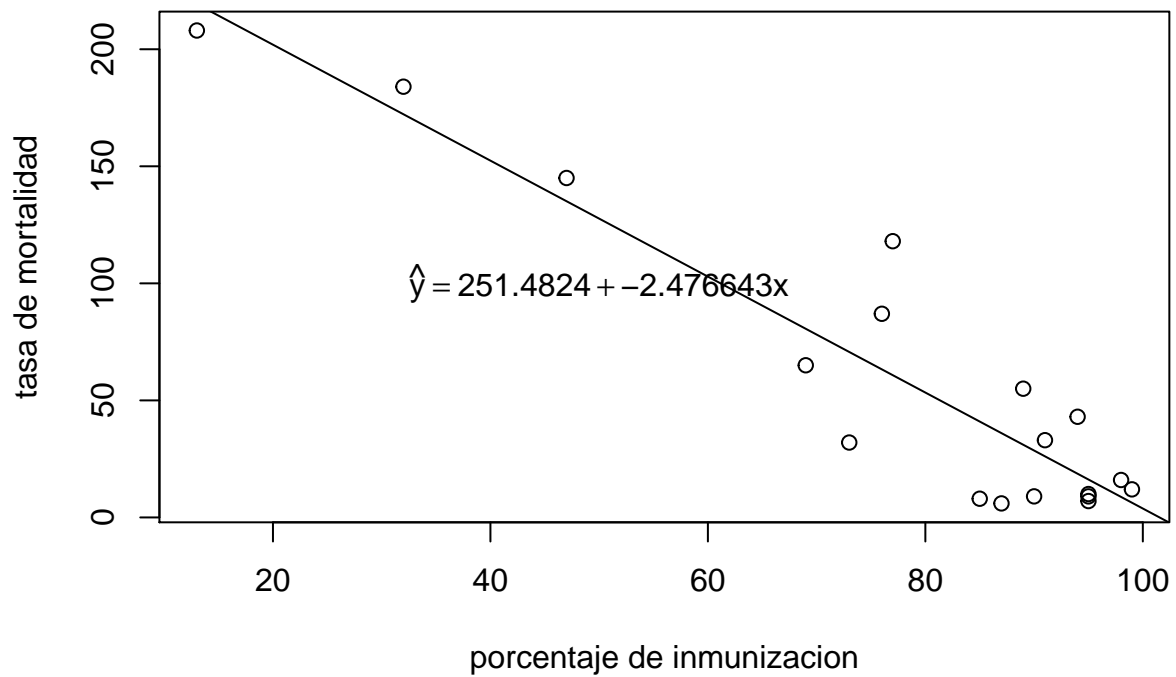
```
porc.inmuniz<-79
porc.inmuniz<-as.data.frame(porc.inmuniz)
predict(13,porc.inmuniz)
```

```
##          1
## 55.58261
```

Extrayendo las observaciones anormales 11 y 12 y creando un nuevo conjunto muertes1 para hacer nuevamenete la regresion

```
muertes1<-muertes[-c(11,12),]
#Haciendo el ajuste por minimos cuadrados excluyendo las
#observaciones anormales y ploteando la linea de regresion para el nuevo
#conjunto de datos
x1=muertes1$porc.inmuniz
y1=muertes1$tasa.mort
l2<-lsfit(x1,y1)
win.graph()
plot(x1,y1,xlab="porcentaje de inmunizacion", ylab="tasa de mortalidad")
```

```
abline(l2)
alfa1=l2$coeff[1]
beta1=l2$coeff[2]
text(50,100,bquote(hat(y)==.(alfa1)+.(beta1)*x))
```



```
ls.print(l2)
```

```
## Residual Standard Error=24.73
## R-Square=0.8617
## F-statistic (df=1, 16)=99.7027
## p-value=0
##
##      Estimate Std.Err t-value Pr(>|t|)
## Intercept 251.4824 20.2188 12.4380      0
## X         -2.4766  0.2480 -9.9851      0
```